## Results Comparison

This study aims to investigate whether increasing the type of pollutants can improve the accuracy of PM2.5 concentration predictions. Table 1 presents a comparison of the results from two modelling approaches: one utilising multiple pollutants and the other relying solely on PM2.5 as the predictive variable.

Table 1 Model performance comparison

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Method** | **Prediction** | **RMSE** | **MAE** | **MAPE(%)** | **R²** |
| Multiple pollutants | Next hour | 0.0318 | 0.0215 | 27.99 | 0.8296 |
| Single pollutant | 0.0334 | 0.0220 | 27.69 | 0.8020 |
| Multiple pollutants | Next day | 0.0330 | 0.0247 | 39.21 | 0.8225 |
| Single pollutant | 0.0349 | 0.0258 | 40.09 | 0.7889 |

For short-term prediction within the next hour, the results indicate that using multiple pollutants did not significantly enhance the model performance. Specifically, the difference in MAPE between the two models was minimal in the short-term prediction. This may be because the impact of other pollutants is less pronounced over such a short timescale. The focus on key features (PM2.5 and meteorological data) in short-term scenarios is sufficient to capture the variations in PM2.5. Therefore, the addition of more pollutants did not result in a noticeable improvement in predictive performance.

In contrast, in the long-term predictions over the next day, using multiple pollutants significantly improved the accuracy. Unlike short-term predictions, long-term forecasts require the model to capture more complex temporal dependencies and interactions between features. The experimental results show that the MAPE and R² of the multi-pollutant model outperformed those of the single-pollutant model. This finding suggests that for long-term predictions, including a broader range of pollutants allows the model to better learn complex dependencies, thereby enhancing the accuracy of the predictions.

In conclusion, while the single-pollutant model is adequate for short-term predictions, the inclusion of additional relevant pollutants is advantageous for long-term forecasting. This finding could assist environmental authorities in formulating more effective policies.